

Figure 1

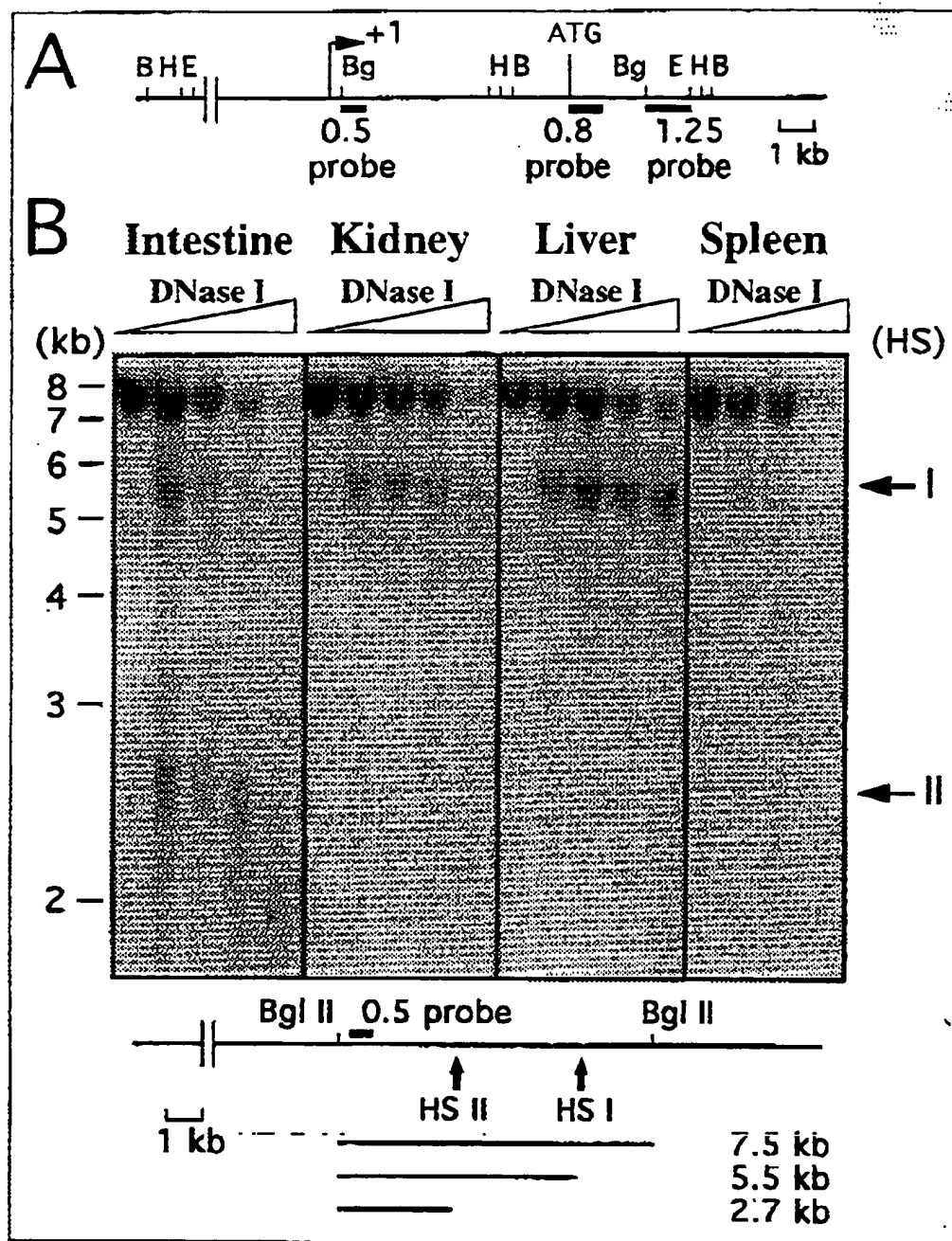


Figure 2. (a)

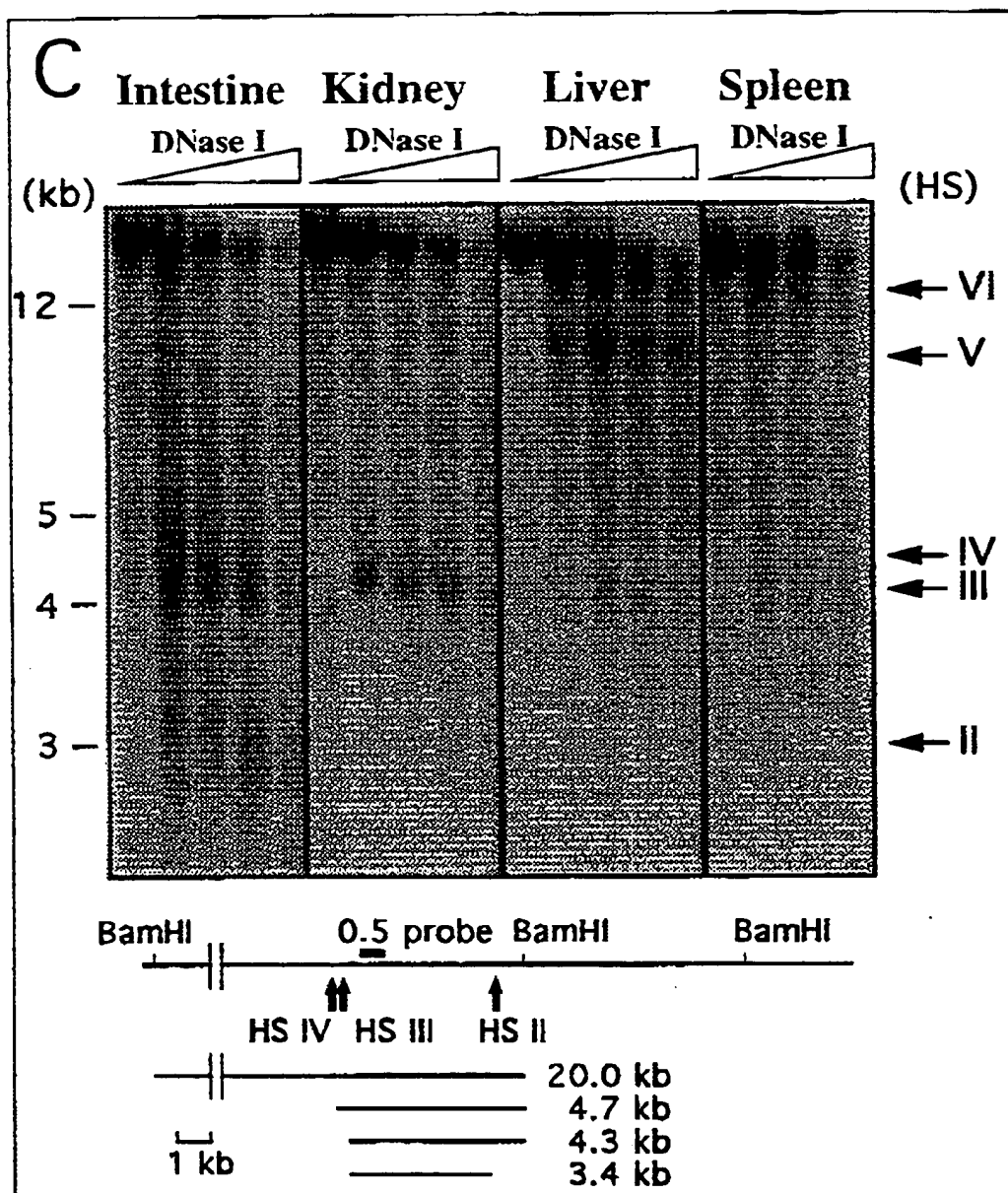


Figure 2(b)

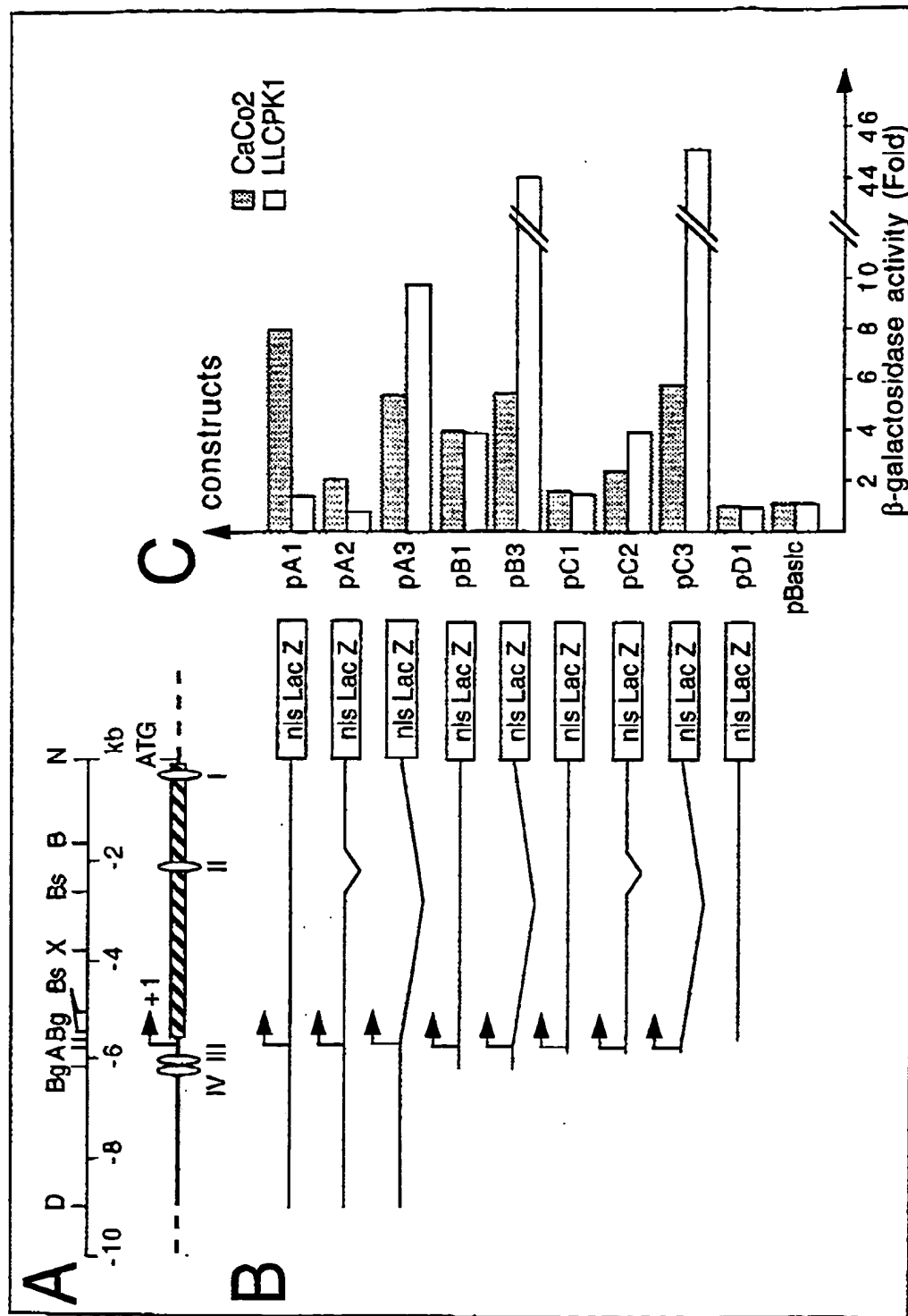


Figure 3

FIG. 4

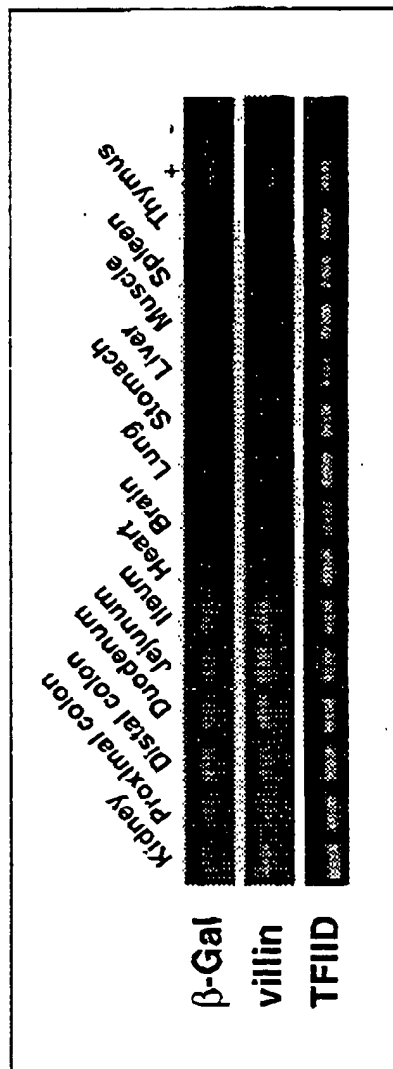


Figure 4



## Genomic sequence of the mouse villin gene regulatory sequences

GATCTGGTGC ACCAAGGACA CTGTGGTCCC AGCACTGGGG AGGTGGAGGG AGGAGGGTCA 60  
 GAAGTTTAAG GTCATCCTTG GTTACATAGC AAGGTTTCAG CCAGCTTCAG CTACATGAAA 120  
 CCTTTGTTTG TTTGTTTGTT TGTTTTAAAG CATTAATAAA TAATACCATA AGGAGGTTGG 180  
 CAGTGGTGGC AGACACCTTT AATTCAGTA TTCAGGAGGC AGAAGCAGGC AGATCTCTGT 240  
 GAGTTCGAAG TCAGCCTAGT CTGCAAAGCT AGTTCAGGA TGGCAAGGGC TACACAGAGA 300  
 AACCTTGTCT CATAAAACCA AAGTAGTAGT AGTAGTAGTA ATGCCATAGA GAAATTTGGA 360  
 GTCCATTTCAG GATGGACCAT CCTATAAGAT GATTCTCTTG ACCCAGGTAA GCTAATGTCA 420  
 TGGGGAAAGG GGATGGGACT GTCCTAGATT AAAAAGTGCT GAGGCGATGC CTATTCTCAA 480  
 TTTGATTCCA TATGAAAAGG CTGATAAGGC CCAAGAGAAG TGGAAGTGGG ACTCTGGACT 540  
 GAAGACGTGA CGGCCTTATA AACACTGGCA CTTATAAACA CTTATAAACA CTGGCACAGG 600  
 CGTTCAGGTT TGAAGATCAC TTTCAAACCA CAGAACAGAA AGTGCTCGCT CGTCCTCAGC 660  
 GTAGCGAGCA CTGGCTGCAG AAGAGTGATA TTTAGTGAAA GCTACCTTCA CAATATCTTT 720  
 GCACCTATCA CATAACGTCG TCAATGTGTC TAACTCCCTA GTCCACAGAT GGCTGTTACA 780  
 CTCGTTTCTG CTTTCCCATC TGGTTGACAT TTGTCAGAAC CAGAAATTAG AAATGTGGGT 840  
 ATTTATTTGT GTGCTGAGGA CACCATCCAG GGCTTTTCAC ATTTCAGGCA CATGGTTTAC 900  
 TAACTGGGCT ACTTCTCCAA CGGTTTGAAA CCATTTGTTT TATATTTACT TATTTTGTGT 960  
 GCATGAGGTA GGCATGTATA CGTATGTATA GGAGTCATGC ATGTGGCTGC TACCTCAAA 1020  
 ATCATTGCAG ATCCCCAGCA AGTGAAGTCA CCGAGCGTTG TAAGTTGTTA TGTGGGACTG 1080  
 GGAGCCAAGG CTGGGTTCTC TGCAAGAGCA GCCAGTGGCC TTAACCATGG GACCAGCTCT 1140  
 CTAGGCCCTAA GGTAATCTTT AGTTTTTTAA AAATATATAT TCTCAGCCGG GTGTGGTGGC 1200  
 ACACGCCTTT AATCCCAGCA CTTGAGAGGC TGAGGTGTAG GAATTATACA CACAGGCCAG 1260  
 CTGGGGTGCA GAGCTTGGCC CTGTTTTTTT TGTTTTTTCT TTATGTGCAC TGGTGTCTTA 1320  
 CCTGCGTGTA TGTCCGTGCA AGGGTGTGAG ATCCCTTGGA GCTGGAGTTA AAGACAGTTG 1380  
 TGATCACGCT GCCGTTACAG ATGCTGGAAA TTGAACCCAG GTGTCCCTAG AGAAGCAGCC 1440  
 AGTGCTCTTA ACTTCTGAGC CACCCCTCCA ACCCTGCTTT TAGAGACTCT TAACCTTTTG 1500  
 TGTAATGTGG GAACTGAGTG GATCTTGCAC TTACCAAGTG TGTGCTGCGC TGTAGCATCA 1560  
 CTGAGCCCGT ACCCACACGA CTAGTGGATA CAGTTTAAGG GCAAACACTT AACAAATGACA 1620  
 ATAGTTGAT AGAGTTTGAA TATAGTCCTG AGCTATTGGT TAGCGTGACC TTTGCTGTCC 1680  
 TTAGCATGTG CTGTGAGAAG ATAGAAAAAT GAAGACTTGA GTCTAGTCCT GGAACCCACA 1740  
 GAGGCAGGCG AGAACCCACT CCTGAAAGTT GTTCTCTGAG CTTACATAC AACTTCACAT 1800

AATAGTTACA ATGATAATAA TAATTAGTAA ATTCTTTTAA AAGGTATATG TTGGGAGGGA 1860  
GAGATGGCTC AGCTTCCAGG AGCACTTGCT GCTCTTGCGAG AGGACCTAGA TTCAGTTCCC 1920  
AGGACTCATA TGGTGGCTCA CAGCCATCTG TAAATCCAGT TCCAGAGGGT TCCACACCCT 1980  
CTTCTGGCCT CCACAGGCAC CACATACATA GTACACAGAC ATACATGCAG GCAAAACACC 2040  
CATACACACA TAAATAAATA AGGAAACTTA AAAGGTGCAT GTGTTGGTAA ACATTGTGCT 2100  
TACACATGCT GATTGAAGAC ATGTACAACG CACACACTGA AGAGGGATCT GGGGCTGGAG 2160  
AGATGGCTCA GCGGTTAAGA GCACTGACTG CTCTTCCGAA GGAAGGTCCT GAGTTCAAAT 2220  
CCTAGCAACC ACATGGTGGC TCACAACCAT CCATAATGAG ATCTGACACC CTCTTCTGGT 2280  
GCATCTGAAG ACAGCTGCAG AGCTACAGTG TACTTAGATA TACTAATAAA TAAATCTTTT 2340  
TTTAAAAAAA TGAAGAGGGA TCTGAGACAC CTCAAAAGAG ATTATGAGCA GTGACTCAGC 2400  
GGTGATTATC TATCCTGGAG TTTTTCCTTT CCGCTTGGCT TGCAACTGGG TGGACAGACG 2460  
CCCCTTTTCA TTCACAAGAA CGGGTGCTAC ATTATTTCTG AACAAAACAG CACCTGCAGT 2520  
ATGTTTACTG TCCTTGCTGA CTATGAGCAC GCGCACGCGC GCGCGCACAC ACACACACAC 2580  
ACACACACAC ACACACACAC ACACACACAC ATTCAGTCTC CAGAGCTCTT GGGGAAGGTCA 2640  
AGAAGAGGCT GCCCTCAAAC ACGATCTTCA TCTTTCCCTC CTAAAGGAGA CCACGATTCC 2700  
AAGGTGGCAG AAGATCTACA GGGGGCAGAG GCAGGGAGGG GGAAGCAGGC CATGGTTTCC 2760  
AGAGACCTAC AGCAGAGGGC AGCAAGGCAG ATCCCCAGGT CCAGGGCAGG GAGGTGGAGG 2820  
CCCTTGTTCC GAGGAGAAGG CAGGCGGCAG AACAGGGTTC AAAGGCACAG GTTTATGGCA 2880  
GCTCATAAAA GTGGAGGTG TGGCTCACTC AGAAAGGAGG AAGAAGGGAA AGGCCCTTGT 2940  
GCCCACTGAG CGAGGGTCAT GCTGAGTAGG AGAGATCTGC AGGGGTGCCA GGAGCCCCAC 3000  
CTGTCTGTCC CAAGGGAACC CCAAGTGTGA ACTCTGGCCT TGGGTGCTGA GTTCCAGCTA 3060  
CAAGACCCCA GGAGTCCTAC TCCATCCCCA TCCAGTGCCC CCTCGCCCCG CCACACCCCA 3120  
CCCCCGACTC CCGTGCCACT TCTCTAGGGC TGGAGGGTGG CCAGCCCTGG TGGGGGTTGC 3180  
CTACCTGCAG GTAGAGCCCA GGTCCTAGCC GGAAGTGCAC CCCATCCCTG AAGCTGCAGA 3240  
GCCAAGGGCG GGGCACACGG CAGCTCAGGC TGTCAGGCTG TTGCTGGGCT CTAGGTTCCC 3300  
AGGGACCTGG GCACCTACTT CCCCACCCCC CCATCCATTC TCTCTGGGGC CCTATCTTCC 3360  
CTTATATGGT GAAGGAAGTT CCTGGGGGGG GGGGGTGGTG GTGAGGACAA AGGTCGTTCC 3420  
GTCTCCTGCA GCCAGCTTGC CACAACTTCC TAAGATCTCC CAGGTGGTGG CTGCCTCTTC 3480

+1

exon 1

(transcription start site)

CAGACAGGTA AGGCAATTGG GTGGGGACAC ATGGTGACCA CAGGTGGTTG GAGGGGACAG 3540  
GGTCCTTGCT TCTCTCTGGC AGCCTGTGCT TTCTGTAGCA CCTTGGTATA AGTTTGGGGG 3600



9 / 15

[illegible]

10 / 15

TCTGGTAGAC CTCTGCTCCT AACTCACCAA GGTATGGCCC ACATTCCTCA CCCAGAAGAG 5580  
 TGCAGAAGAG AGCCTTAGAG AAAGGGTAAC AGTAACAAAG ATGGCCAGAA TAAAACAAA 5640  
 ACTACTATCC TTTGTACCCA AATTGGTTTT GCTGAACCAG GAGGGGGTGT GTGAGTGTAT 5700  
 GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT 5760  
 CTTGGGGGAC TTTTCATGCT AAAGAATATC TGATATTGGC GCCCATGCCA ACAGGGGTAT 5820  
 TGGGGAGAGT CAGGCTTCTG CAAACACAGT AAGCTGCCCC AGATGGATTG GTGGCCTGAA 5880  
 TCACCAAGGG GCAGGCTGAT CAGAGTGGAC AGAACATCAC AAGATAAGCC ACCCTGTGGG 5940  
 GCTCAGAAGA GGGAGTTTAC AAGAGGTAAA GGCCAAGCCA TTTATTATCC AAGACATGAC 6000  
 TCAAAATCAA AGTGCAAGGA GAGATTAGCT GGAGAGATGG GGCTGTCAGT GTGGGACACC 6060  
 TGACCTTGCA CTTATTAGTC ACTAGGCCAA GGAGCAGTCA CAGAGGGTGA CTGGGTCCTA 6120  
 CTCAGCTTGG AGCAGGCACG TGGAGAATGG GTGACCTCCA TCCTGATGGA GAGGGCTGAG 6180  
 CACCACCAGG TACAAGTGTT CCCTGTGTCT CATGCCAGGA TTCCTGGCCA GTTTTCAAAG 6240  
 GACTAAGGAC TCATCTCTGG TGGAAACAAA GTATCCAAGC CCTAAGCCCC ATTTTGGTCT 6300  
 AATTAAATCA GAACCCCTGG GGATGCAGGC TCTGAGCAGC AGGAGCTTTT TAAAAAGCTC 6360  
 CCAGGTGATT CTGATCAGCA GCTGGAACAA ACACAGCTAC AGGTTCAAAC AGAAAGAGGC 6420  
 AAAGCTAGGG AAAGCTTGGG ATGGGGAGCC TTCTTCCAGG CCAGTAGATG GAGGCTGGTT 6480  
 AGCAGTGGTG GCAGCTTCTC TCTGCCTGTC ATATAGCTAT CCATCCACTC ATCCATCCAT 6540  
 ACACCCACCC ATCCATTTAT GCACCCATCC TTCCATCCAT CCATCTATCC AGCTACCCAC 6600  
 CCACGCATCC ATCCAAACCT TCCTTTTCTC CTTCTTTCTT TCTTTTTTCC TTCATCATT 6660  
 CATTTATCCA ACAGAGAACT GGTATTGTAC TAAATGTGGG AGATTTAATT AATTTTAGA 6720  
 AGCTCTGTTG ATTGACTGAT TGTGCATGTA TGTGGACAGG TACATAACCAC AGCACACGTG 6780  
 TGGCAATCGG AGAAAGGTTT TGGGTGTTGT TTTCTCTTCC CACCGTGTGG GTTCTGGGGA 6840  
 TTGAACTCAA ATTATCGGGC TGGTGGCAAG TGTCTTTACC ACCGAGCCAT TTTGCTGACA 6900  
 CATCATTATT ATTAGAAAGC ATCTTATGTA GTCCAGGCTG GCCTCAAGCT TGCTATGTCG 6960  
 CCACGGATGA CCTTTAACTC CTGCTCTTCC AGCCTCCACC CGAGTGCTAG GTTTACAGGT 7020  
 GTTCAACTGG TGAATGCCTT TAATCCCAGC ACTCTGTGGG GGGGGGGGGG GAGGCGGATC 7080  
 CCTGAGTTGG AGGCCAGTTT GGTCTACAGA GTTTCAGGAT ACCTGGGGCT ATACAGGGAA 7140  
 ACCCTATCCC AAACAAACAA ACAACAAAC AAAAAATATT CTGTGCAATA ATCACAGAGA 7200  
 TTAGAGGATA TTAGTAGGGT AGTAGGGCTG GTGAGGGAGA GTCATGCTTT CTTTTGTATT 7260  
 ATAATAGTAA AGTACTCACA AGATGCATTA TCTATCTATC TATCTATCTA TCTATCTATC 7320  
 TATCTATCTA TCTACCTACC TACCTACCTA TCCATCCATC CATCTATCGT ATAGCCCAGG 7380  
 CTGCTTTGAC TCTGAATGCT CCTATTCTG GGTCAACTCT TCACCCCTAG TGTTGGGTTT 7440

11 / 15

ACCAACACCC AGACATTTAT TTTATTTTGT TTTATTTTAT TAATCTAGGA GCTCAGGGTG 7500  
 GGACTCAGGG TCTTGTGCAT GCTAAGCAAG CTCTCTGCCA CAGAGCTGCA GCTCCAGTCC 7560  
 CCATTTTGTG CAGGTGACTC TGTGACAGTT GTCATATTCTG CAGCGCTATG TAGCTCTCTC 7620  
 CACCTCCCAG TTCCAGCACT TTCTGGTCAT CCCAGTGGGC GGGCAACTCT GTGCTCACCA 7680  
 GTGCCCTGTT CCCGTCTTTC AGACCTACAT ATTTGCCTGT CTGAACAGTT CATGTAAATG 7740  
 GGATGCGTTC CTGTGTATTTC TTTTATGGCT GGCCCCTTTA TCTTAGCACA GTTTGTGTTG 7800  
 GGCCATGTGT CACTGCTATA CTCTATCTTA TCATCATCTT ATGGCTTAAT AGTGTTCCTT 7860  
 TGTGTGGATA AACCACCTTC TGTTCATTT ACTGATGGAA ATTTGTGGCC CCACCCCCAC 7920  
 CCTTTTTTTT TTTATTTGAG ACAAGGTCTT TCTGTGTAAT CTTGCAATCT TGGCTGTCCT 7980  
 GAGCTCACTC TGTAGACCAG GCTGTGAGGC TGTCTTCCA CTTTGTGACAC TCCTGTGAAC 8040  
 AGAGTAGCCA TGAACCTCAA AGACAATTTT CTGTTTTGGT TTGTTTTTTA CATTTGTGTG 8100  
 TGTATGCGTG TATATGTGCA TGTTTGTGTC TTCAGGTGCT CACATGTGTG TACCTGTGTG 8160  
 TGGGACAGAG AACAAACCGA TGTGCCATTC CTCAGATACT ACGCATCTTG TTAATATGTA 8220  
 TGTATTATGT ATGTTTATTT AGTGTGCCCA AGTATGCAGG TATTTTGTG GAGTTTTTAC 8280  
 CTTCCCTTGT GGGCTCTCCG CATTAAGTCTC AGCTCCTCGG GCTAGTGAGC AATGCCTTCA 8340  
 CTCGATGAGC CATCTCGCTG CCCCTGCTGC CACCTCCTCC TTATTTCCCA GATGGGACTA 8400  
 CGCACTGCAC TGGCCTAAAG CTCACCAAGT CATCCAGAGT GGCTAGCCAG GGAGACTCAG 8460  
 GGATATGCTG GCCTCTGCCT CCACAGTGCT AGAATTACAG GCATACATCA CTGCTGGAAG 8520  
 ATTTTAAACC TGAATCCTGA GGATAGAGCA GGCAGTCTAC CAATGGAGGG TTCTTTTGT 8580  
 GTTTGGTTTG GTTTCCTCTG CATAAGATCA GGCAGTCTGA AATAGTGTAG CCTGGGCTAC 8640  
 ATAACATCTT GTCTCAAAAA GCCTATAGAG GTAGGGAGGT CGAGGCTAAA GAAGAGCCTT 8700  
 AAGCCGGCTG TGATAGCACA CAGGATAGCC TGCATATAT AGCAAGACCT TGTTCAAAA 8760  
 ACATGGAGGG AGGGGTATGT TTTAAGTGCT GGGCTGTGTA ACAGGCACTA AGGGAGCCAA 8820  
 TGTAGACATT TGAATAAGAA AGGATCATCA TCAAAGCCGG GTGGGCAGGG TAGAGGTTGG 8880  
 ACTACAGTGG TCAAGACCCC CATAGGAAGC CAGTTTCCCT TCTTCCTCTG GGCTCAAGC 8940  
 CTGGCTCGAC GGCCACTGCT CTCACATGCC TTCTCCTCTA GGCTCGTCCA CCATG 8995

exon 2

# Targeting using the regulatory sequences of the mouse villin gene

Data obtained by transgenesis - October 1998

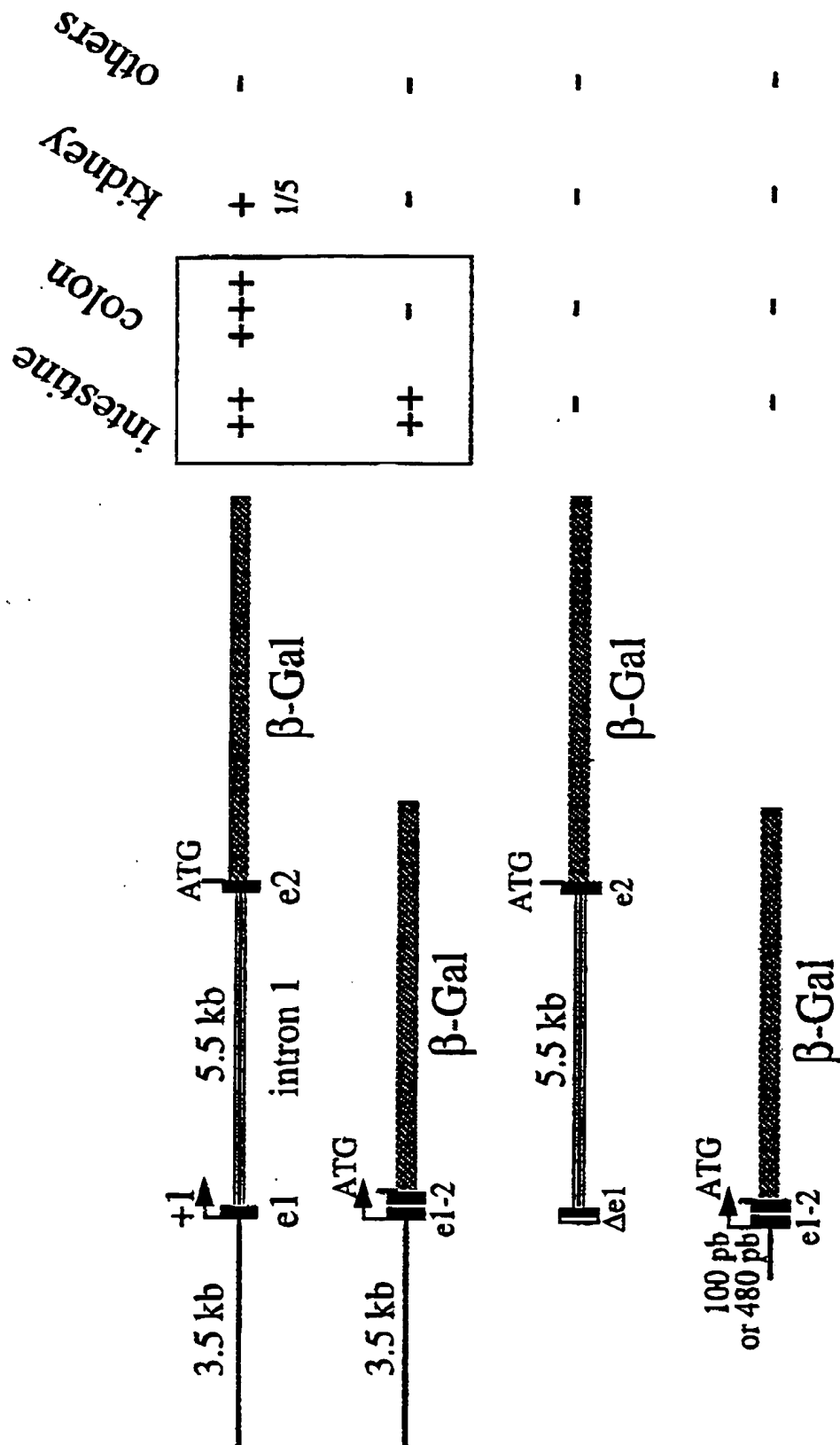


Figure 7

### Targeting of oncogenes and tumor suppressor genes

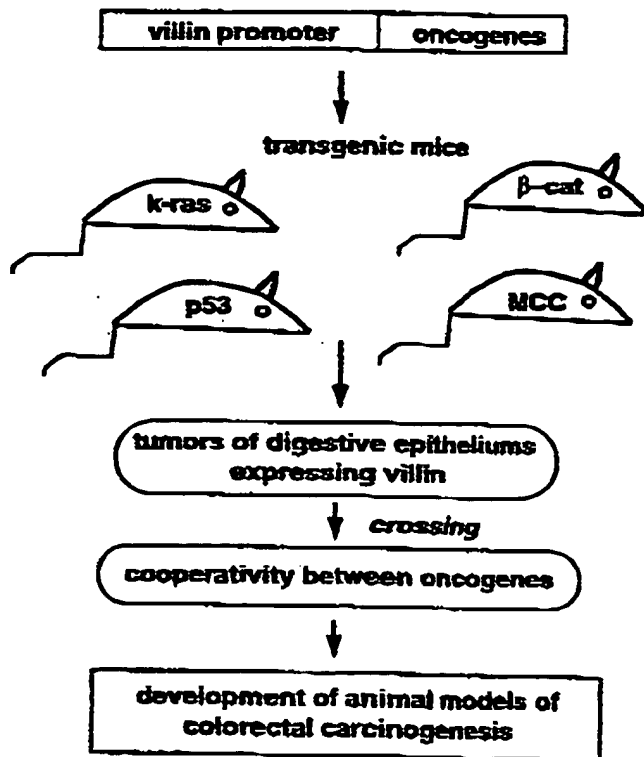


FIGURE 8A

### Targeting of immortalizing gene

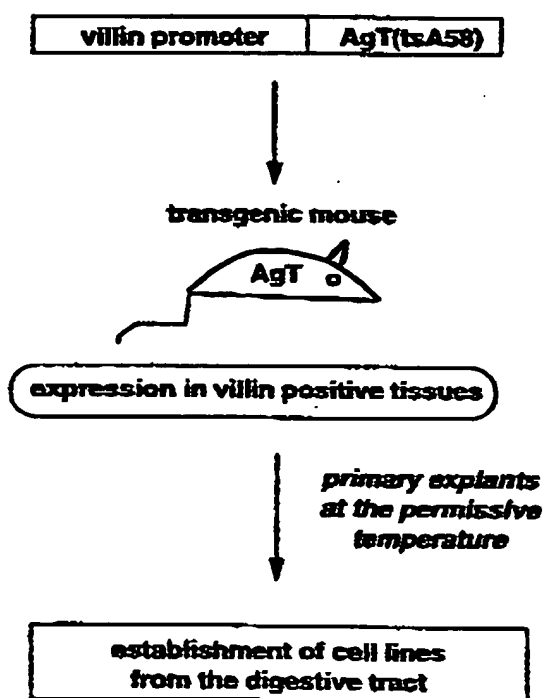


FIGURE 8B

15 / 15

### Targeting of transactivator gene (repressor form rtTA)

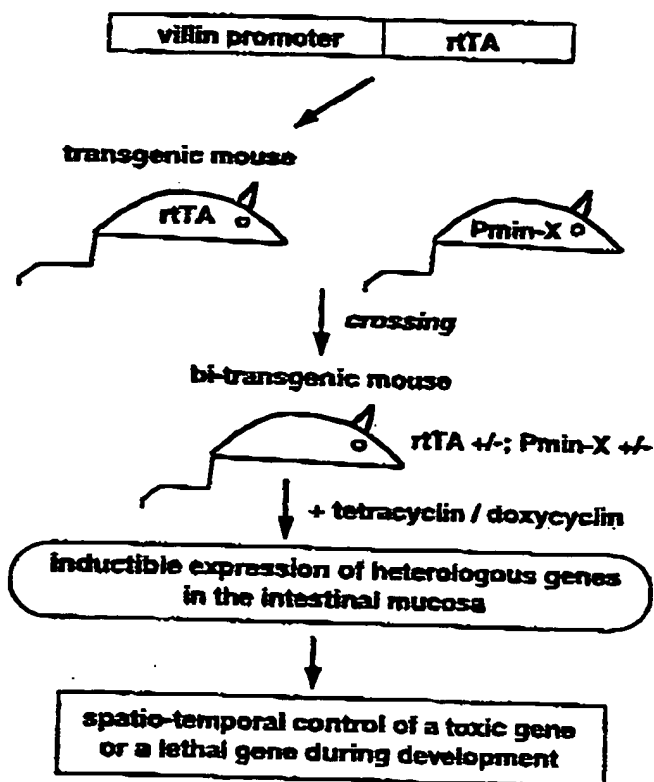


FIGURE 8C